Non-Agentic Civic AI - Illuminating America's Civic Digital Field: Volume II

CIVIC CAPACITIES ALIGNMENT

***** CIVIC PROBLEM-SOLVING and ARCHIVING for the DIGITAL AGE *****

The U.S. Citizenry's Distributed Civic Capacity,
Continuous Interdisciplinary Expert Committees,
and Non-Agentic Civic AI

Christopher Hunt Robertson, M.Ed.

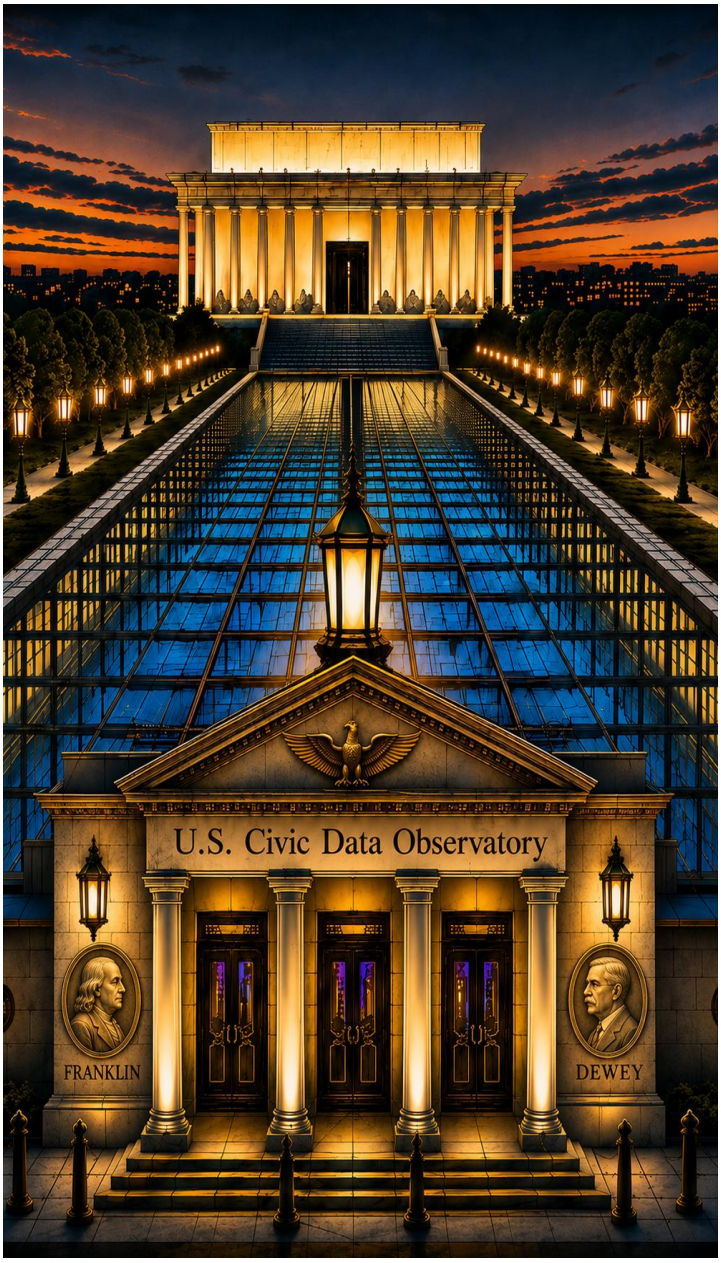
(With the Collaborative Support of Advanced AI - ChatGPT, Perplexity, Gemini/Google)

Honoring the Democratic Vision of
Benjamin Franklin, John Dewey,
and the
Foundational Authority
of the United States Citizenry

Volume I: *Civic Digital Visibility First*: Introduces a visibility-first design principle and offers specific architectural models that envision U.S. Civic AI as strictly public non-agentic infrastructure. Calls for the replacement of Agentic AI with Non-Agentic Civic AI in most U.S. government-related functions at all levels, with some exceptions for military, intelligence, and law enforcement agencies. Proposes Civic AI coordination by the Franklin - Dewey - Citizenry Civic Data Observatory, and nonpartisan supervision by the Library of Congress.

Volume II: *Civic Capacities Alignment:* Envisions the Citizenry's Distributed Civic Capacity, Continuous Interdisciplinary Inquiry Committees, and Non-Agentic Civic AI all cooperating through - but not directed by - the Franklin - Dewey - Citizenry Civic Data Observatory. Also considers how, as the U.S. prepares for its AI-influenced future, concentrated power will need revisiting, with new public programs, new civic tools, and an updated social contract required.

Volume III: *52-Gates Democratic Permissions Certification Threshold*: Identifies the democratic principles which Civic AI must never violate, and proposes a constitutional admissibility standard by which all legitimate U.S. Civic AI systems must be certified. Provides a strict binary refusal framework to enforce democratic principles necessary for continued self-governance.



Civic Capacities Alignment – Theoretical Statement

Civic Capacities Alignment recognizes existing civic capacities and their potential for complementarity while respecting the established roles of democratic and legal institutions and governmental processes. Specifically, Civic Capacities Alignment focuses on supporting the **Citizenry's Voluntary Distributed Civic Capacity**, **Continuous-Inquiry Interdisciplinary Expert Committees**, and **Non-Agentic Civic AI** (NACAI) as complementary civic capacities.

The alignment sought is not an alignment of machine values with civic values, nor an attempt to model, predict, or represent collective democratic conscience. It is a constitutional division of labor in which voluntary citizen authority, accountable continuous inquiry, technological illumination, and lawful institutional action retain distinct roles. Citizens require greater shared visibility into complex civic conditions; interdisciplinary committees require accessible, continuously organized evidence; and NACAI requires human conscience and democratic judgment to give its observations legitimate public meaning.

Within this framework, NACAI serves as **Civic Visibility Public Infrastructure -** an auditable, non-agentic infrastructure that preserves public records, organizes and reveals relevant conditions and aggregate patterns, identifies system-level pressures, and helps institutions and communities recognize emerging conditions earlier. It must not score individuals, recommend policy outcomes, trigger enforcement, or exercise delegated governmental judgment. Judgment, political deliberation, and public responsibility remain with human citizens and institutions.

The theorist proposes a **U.S. Franklin-Dewey-Citizenry Civic Data Observatory** within the **nonpartisan, civilian Library of Congress** as a potential institutional home for coordinating NACAI's civic visibility function. Its role must be archival as well as illuminating: preserving, authenticating, and making intelligible the evidentiary record in a public environment where civic data is reproduced, compressed, recombined, and interpreted at scale. Without provenance-based authentication - the verification of civic data's origin and integrity - democratic judgment cannot remain reliably grounded in evidence.

Core principles of Civic Capacities Alignment, presented in *Non-Agentic Civic AI - Illuminating America's Civic Digital Field*, include No Moral Outsourcing, Digital Nature Integrity, Amoral AI Civic Advantage, and Moral Visibility.

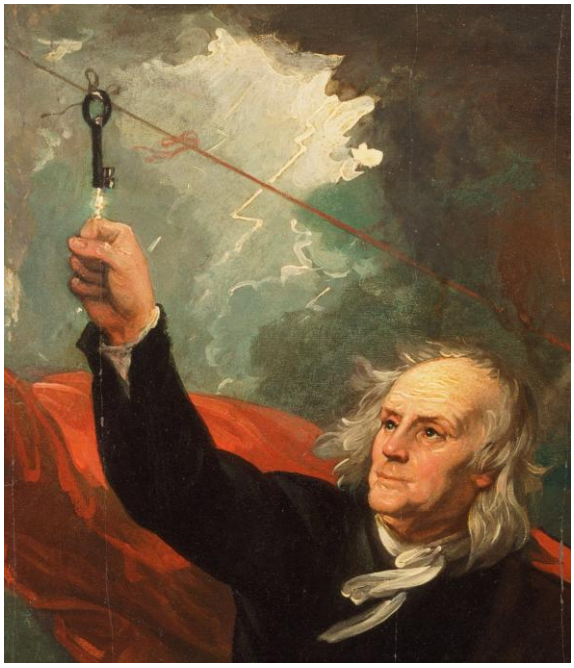
Non-Agentic Civic AI may be highly capable without becoming unaccountable civic power or eroding democratic self-government.

Please Note: This book is being iteratively published in sections in 2025 and 2026. More material may currently be found under the author's name at Medium.com and Academia.edu.

"Draw the Electrical Fire out of a Cloud before it came nigh enough to strike …"

BENJAMIN FRANKLIN (Realist)

Democracy and Defense Advocate, Civic Builder, Scientific Inventor, and a U.S. Founding Father

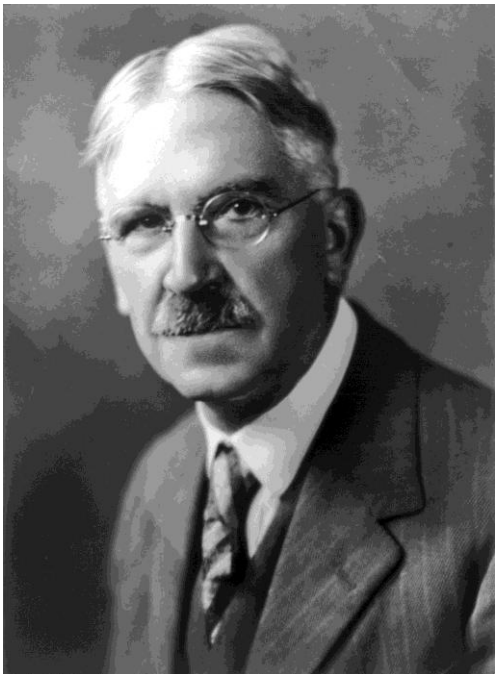


Franklin encouraged the building of civic infra-structure, public service, the preservation of information, and the open dissemination of knowledge to the public. Like Franklin's kite and key experiment, which made invisible forces visible to human understanding, this Civic AI proposal seeks to reveal hidden patterns in civic digital life, make systemic conditions perceptible, and inform public deliberation.

"Democracy has to be born anew each generation and education is its midwife."

JOHN DEWEY (Pragmatist)

Democracy Advocate, Philosopher, and a U.S. Education Reformer



Supporting Dewey's core beliefs, this Civic AI proposal emphasizes that healthy democracies require citizens' systemic awareness, collective problem-solving, and participation in governance. Citizens must actively and routinely shape society through continuous inquiry, education, and open debate. Otherwise, they will become overwhelmed, passive, and disconnected from governance due to the complexity of modern industrialized society - "the eclipse of the public."

2026 State Presentation: Ohio Academy of History Presentation (Marietta College, 2026):

Civic Digital Visibility First: A Conditional Welcome to Our Emerging Non-Agentic Civic AI - Supporting 2500 Years of Democratic Evolution

2027 Regional Presentation: (Scheduling):

Civic Problem Solving: Civic Capacities Alignment and Proposed U.S. Non-Agentic Civic AI Public Infrastructure

"Democracies cannot govern what they cannot see."

Christopher Hunt Robertson, M.Ed.

Three-Volume Arc:

New Technological Infrastructure and Societal Reforms for American Renewal

Volume I Focus: Non-Agentic Civic AI

In the AI-accelerated phase of America's continuing New Deal public-works era, democracies require different "smart tools" than the agentic systems now being deployed—non-agentic civic visibility instruments that illuminate events, structures, patterns, and risks without exercising authority or automating judgment.

Volume II Focus: Preparing for the Epic Democratic and Societal Challenges Ahead in America's Age of Humanity with AI

Long-standing tensions between concentrated economic and corporate power and broad democratic expectations have repeatedly shaped American institutional development, and the emergence of AI appears to be accelerating many of those existing dynamics. As democratic societies adapt, we may require new public institutions and civic tools - such as the *Non-Agentic Civic AI* infrastructure proposed in Volume I and constitutionally grounded in Volume II - to help citizens and their institutions perceive, understand, and navigate our current AI-accelerated industrial transformation.

Beyond developing specialized Civic AI, concentrated economic and corporate power will need to be revisited, much as they were during the New Deal Era. Major societal reforms will likely be needed, including an update of America's social contract. As President Franklin D. Roosevelt, noted during the Great Depression: "*The test of our progress is not whether we add more to the abundance of those who have much; it is whether we provide enough for those who have too little*." - Second Inaugural Address, January 20, 1937.

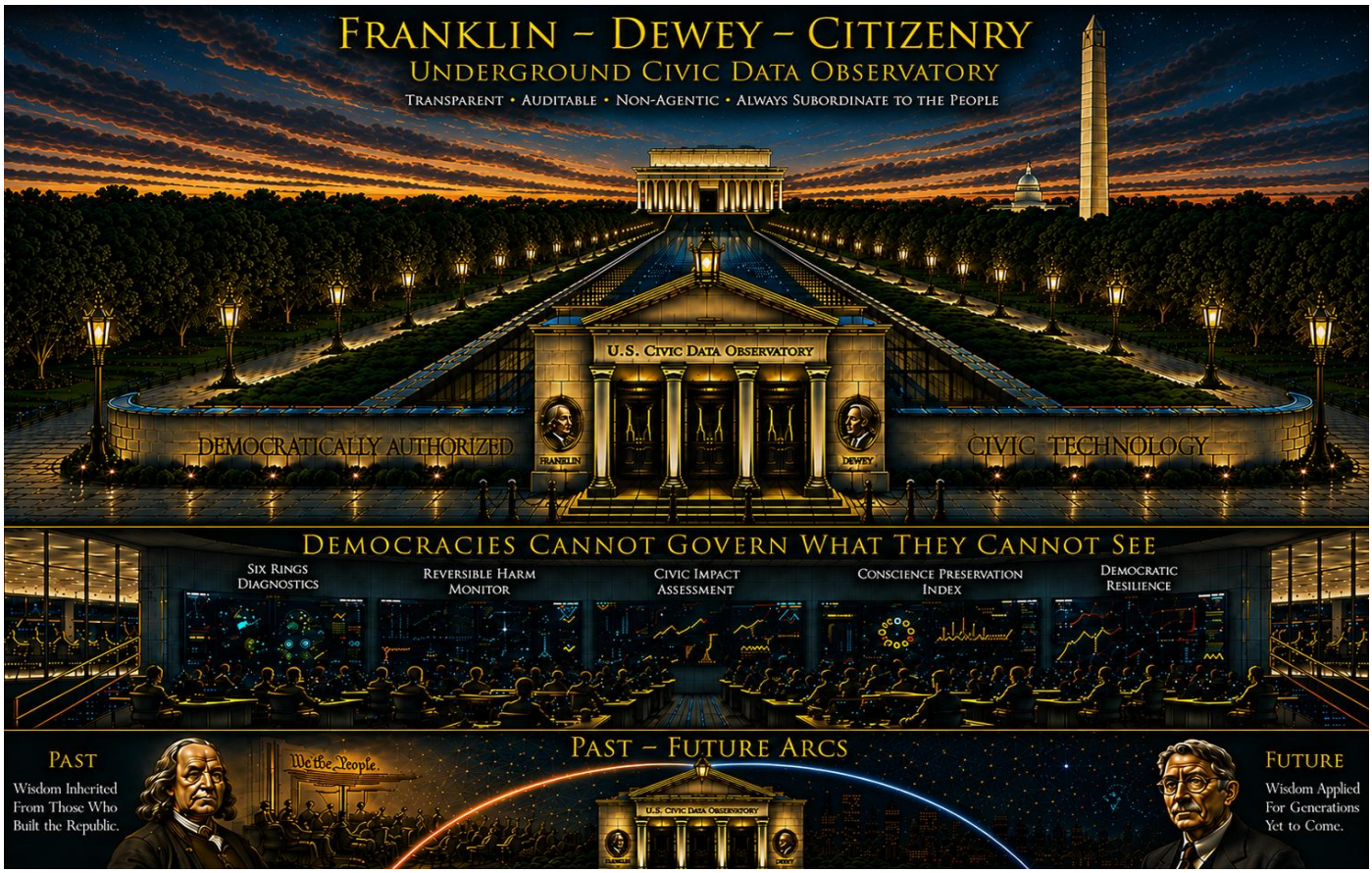
Volume III Focus: Non-Agentic Civic AI's Democratic Foundation: *The 52-Gates of Democratic Permissions Threshold*

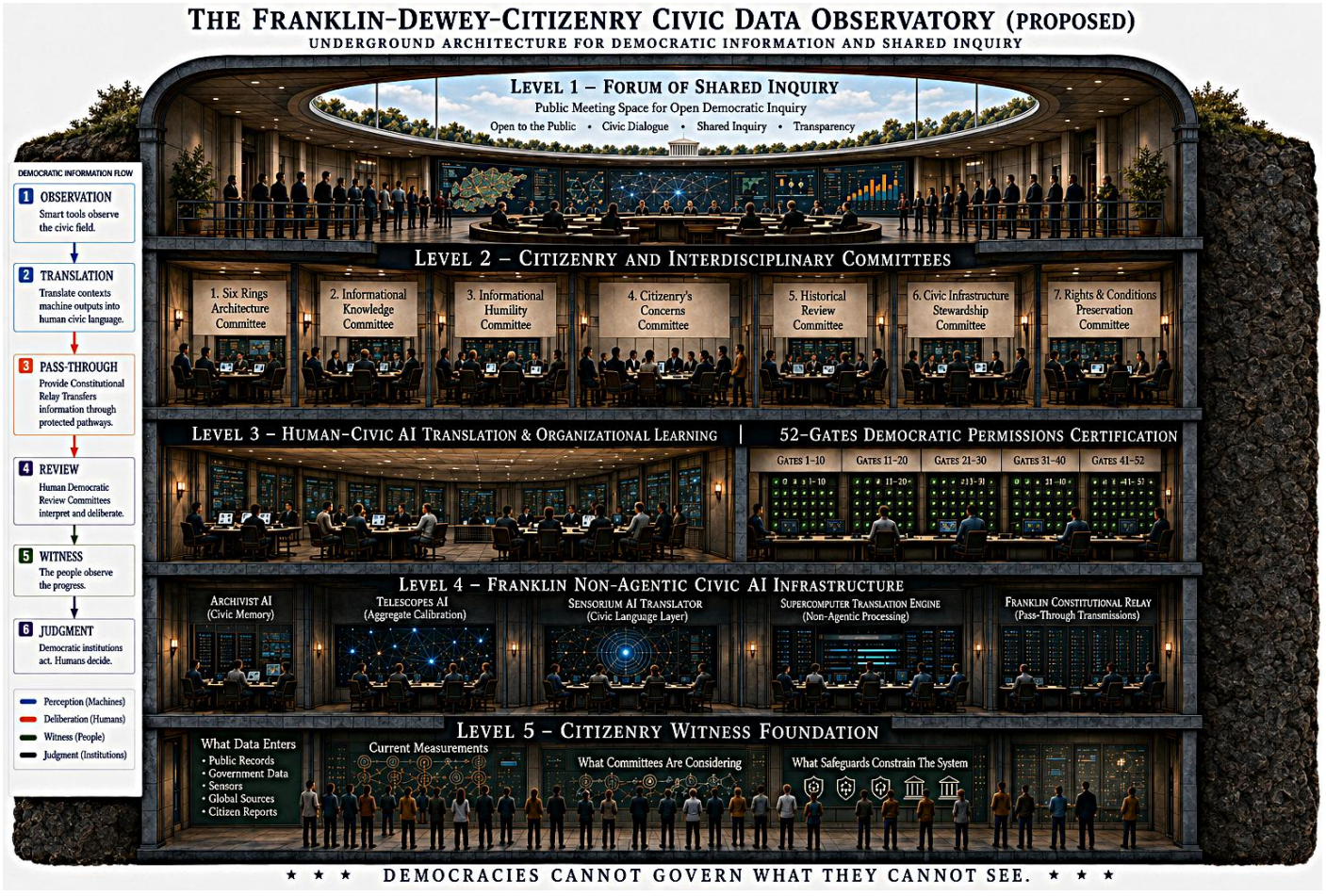
In a constitutional democracy, legitimate authority can arise only from the citizenry and the Constitution, informed by enduring democratic principles developed through historical experience. Non-Agentic Civic AI must therefore remain permanently subordinate, structurally constrained, and constitutionally admissible—never a source of independent power or a substitute for human moral sovereignty.

This book series asserts that the discontinuity we're experiencing is technological, not constitutional. Enduring democratic principles and collective moral responsibilities remain in force, but they now demand renewed civic infrastructure and major governmental reforms to successfully transition our citizenry through our current AI-accelerated industrial transformation.

Please Note: This book is being iteratively published in sections and should be completed within 2026. More material may currently be found under the author's name at Medium.com and Academia.edu.

(Image by ChatGPT)


FRANKLIN – DEWEY – CITIZENRY
UNDERGROUND CIVIC DATA OBSERVATORY
TRANSPARENT • AUDITABLE • NON-AGENTIC • ALWAYS SUBORDINATE TO THE PEOPLE
U.S. CIVIC DATA OBSERVATORY
DEMOCRATICALLY AUTHORIZED
CIVIC TECHNOLOGY
DEMOCRACIES CANNOT GOVERN WHAT THEY CANNOT SEE
SIX RINGS DIAGNOSTICS
REVERSIBLE HARM MONITOR
CIVIC IMPACT ASSESSMENT
CONSCIENCE PRESERVATION INDEX
DEMOCRATIC RESILIENCE
PAST
Wisdom Inherited From Those Who Built the Republic.
We the People.
PAST – FUTURE ARCS
FUTURE
Wisdom Applied For Generations Yet to Come.


THE FRANKLIN-DEWEY-CITIZENRY CIVIC DATA OBSERVATORY (PROPOSED)
UNDERGROUND ARCHITECTURE FOR DEMOCRATIC INFORMATION AND SHARED INQUIRY
LEVEL 1 – FORUM OF SHARED INQUIRY
Public Meeting Space for Open Democratic Inquiry
Open to the Public • Civic Dialogue • Shared Inquiry • Transparency
DEMOCRATIC INFORMATION FLOW
1 OBSERVATION
Smart tools observe the civic field.
2 TRANSLATION
Translate contexts machine outputs into human civic language.
3 PASS-THROUGH
Provide Constitutional Relay Transfers information through protected pathways.
4 REVIEW
Human Democratic Review Committees interpret and deliberate.
5 WITNESS
The people observe the progress.
6 JUDGMENT
Democratic institutions act. Humans decide.
Perception (Machines)
Deliberation (Humans)
Witness (People)
Judgment (Institutions)
LEVEL 2 – CITIZENRY AND INTERDISCIPLINARY COMMITTEES
1. Six Rings Architecture Committee
2. Informational Knowledge Committee
3. Informational Humility Committee
4. Citizenry's Concerns Committee
5. Historical Review Committee
6. Civic Infrastructure Stewardship Committee
7. Rights & Conditions Preservation Committee
LEVEL 3 – HUMAN-CIVIC AI TRANSLATION & ORGANIZATIONAL LEARNING
52-GATES DEMOCRATIC PERMISSIONS CERTIFICATION
GATES 1–10
GATES 11–20
GATES 21–30
GATES 31–40
GATES 41–52
LEVEL 4 – FRANKLIN NON-AGENTIC CIVIC AI INFRASTRUCTURE
ARCHIVIST AI (Civic Memory)
TELESCOPES AI (Aggregate Calibration)
SENSORIUM AI TRANSLATOR (Civic Language Layer)
SUPERCOMPUTER TRANSLATION ENGINE (Non-Agentic Processing)
FRANKLIN CONSTITUTIONAL RELAY (Pass-Through Transmissions)
LEVEL 5 – CITIZENRY WITNESS FOUNDATION
What Data Enters
• Public Records
• Government Data
• Sensors
• Global Sources
• Citizen Reports
Current Measurements
What Committees Are Considering
What Safeguards Constrain The System
★ ★ ★ DEMOCRACIES CANNOT GOVERN WHAT THEY CANNOT SEE. ★ ★ ★

(Images by ChatGPT)

Introducing a New Civic AI Approach …

THE "CIVIC DIGITAL VISIBILITY FIRST" ARCHITECTURAL STACK

"Democracies cannot govern what they cannot see."

The **Civic Digital Visibility First Architectural Stack** treats civic digital visibility as the prerequisite condition of democratic governability. Rather than centering optimization, prediction, or automated decision authority, it proposes visibility-only public-works that restore shared sight and deliberative time under machine-speed conditions.

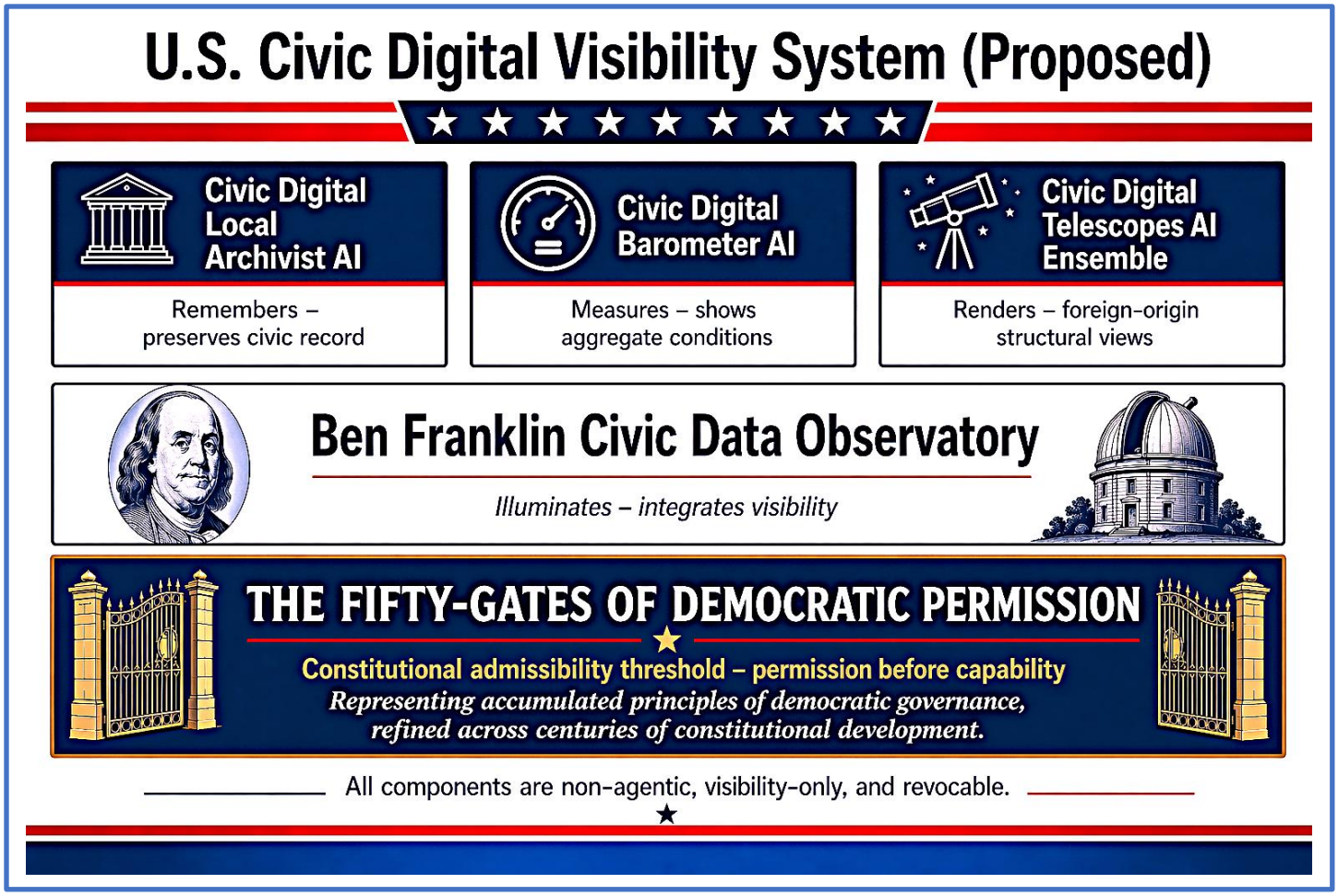
The stack's components - each within current technological reach - address civic digital invisibility, not domestic privacy regulation or citizen surveillance. Inspired by Benjamin Franklin's public-works ethic, the design is non-agentic, rights-tethered, revocable, and refusal-bound.

Six Rings of Digital Invisibility and Civic Destabilization (Problem Diagnosis): The Six Rings provide the diagnostic lens for the entire architecture. They describe how fragmented information flows, opaque digital systems, and accelerating machine-speed processes erode shared civic visibility before overt constitutional crisis appears. By tracing how civic harms accumulate across successive "rings" of invisibility - from overload and attention capture to institutional blindness and destabilization - the model reframes alignment as a problem of shared perception and consequence representation rather than abstract technical control. The remaining components of the stack are designed to restore sight across these rings - without transferring authority from humans to machines.

The Ben Franklin Civic Data Observatory (Border-Facing National Visibility Infrastructure): The proposed Ben Franklin Civic Data Observatory is the border-facing visibility instrument of the stack. It is a purpose-built, non-agentic civic infrastructure designed to detect foreign-origin digital interference at aggregate and system levels. Modeled on Franklin's public works - lamps, levees, and lighthouses - the Observatory extends shared sight across institutions while leaving judgment and coercive authority entirely in human hands. It does not conduct domestic surveillance. It does not support enforcement. It does not classify persons. It detects foreign-origin systemic patterns and refers findings to existing constitutional authorities, who remain free to act - or not - with eyes open. The Observatory is admissible only under the 52-Gates Democratic Permissions Threshold and remains revocable, suspendable, and dismantlable by democratic decision.

52-Gates Democratic Permissions Threshold (Constitutional Admissibility for All Civic AI): The 52-Gates of Democratic Permissions is a pass/fail constitutional instrument for admitting artificial intelligence systems near democratic authority. Unlike flexible, risk-based governance frameworks, the 52-Gates ask a prior question: under what structural conditions, if any, may machine systems operate near democratic power at all? Each Gate articulates a democratic principle, a structural necessity, and a binary disqualifying test. Systems that fail any gate are inadmissible. The 52-Gates instrument prohibits delegation of moral judgment to machines, forbids enforcement or optimization authority, requires public visibility of civic-scale outputs, and embeds refusal, revocability, and periodic re-certification as ordinary safeguards. No system operating near democratic authority may qualify as constitutionally legitimate without passing every gate.

Non-Advisory Civic AI Archivists (Domestic Democratic Illumination AI) and **Democratic Civic Impact Assessments** (Their Analytical Reports)**:** On the domestic side of the constitutional line, Non-Advisory Civic AI Archivists are strictly non-agentic, constitutionally bounded intelligent tools designed to preserve democratic visibility and institutional memory. The Archivist maintains structured civic records; retrieves and cross-references laws, debates, and prior impact statements; and, when explicitly authorized in declared Analytical Mode, generates Democratic Civic Impact Assessments, which are machine-speed, non-binding descriptive analyses of projected or observed civic-scale consequences. These assessments may map burdens and benefits across groups and time horizons, compare scenarios using predefined metrics, and surface structural tensions among civic goods. They may never recommend, rank, optimize, aggregate into composite civic scores, determine legitimacy, or trigger action. Machines may illuminate trade-offs, but humans must judge them.


U.S. Civic Digital Visibility System (Proposed)
Civic Digital Local Archivist AI
Remembers – preserves civic record
Civic Digital Barometer AI
Measures – shows aggregate conditions
Civic Digital Telescopes AI Ensemble
Renders – foreign-origin structural views
Ben Franklin Civic Data Observatory
Illuminates – integrates visibility
THE FIFTY-GATES OF DEMOCRATIC PERMISSION
Constitutional admissibility threshold – permission before capability
Representing accumulated principles of democratic governance, refined across centuries of constitutional development.
All components are non-agentic, visibility-only, and revocable.

NOTE ON AUTHORSHIP AND THE USE OF INTELLIGENT TOOLS

This book was written by Christopher Hunt Robertson, M.Ed., with the support of advanced intelligent tools. Two generative large language models, ChatGPT and Perplexity, assisted with earlier drafts, while two Google large language models, Gemini and Google Search AI Mode, provided reviews and suggestions for refinement. These systems contributed substantially to brainstorming, research assistance, literature synthesis, comparative reasoning, logical stress-testing, and writing refinement. ChatGPT helped to skillfully create the books' images.

Two additional large language models, Claude and Copilot, occasionally provided supplementary information, alternative perspectives, and further stress-testing of the framework introduced herein.

Throughout the project, these systems were treated as non-agentic intellectual instruments: tools that assist human thinking without originating judgment, authorship, or moral responsibility. Their capacity to contribute toward civic ends was appreciated.

The human author directed, curated, revised, and ultimately judged all arguments and claims. He assumes full responsibility for the content of this work, including any errors or omissions. Prior to the availability of such tools, the author was the recipient of nine book awards from historical associations.

At the time of this writing, there remains active scholarly and public debate concerning the appropriate role and acknowledgment of AI assistance in intellectual work. This project relied on a simple premise: that the search for truth is best served by transparency about methods, restraint in the use of tools, and clear retention of human responsibility for judgment.

Non-Agentic Civic A.I.: Illuminating America's Civic Digital Field (Vol. I)

CIVIC DIGITAL VISIBILITY FIRST

Continuous Civic Digital Resolution for Democratic Restoration

Table of Contents: (To Be Uploaded Soon)

Part X: CONTINUOUS CIVIC LEARNING and RESOLUTION: BRIDGING COLLECTIVE CONSCIENCE to DEMOCRATIC RESPONSIBILITY (Inspired by John Dewey's Pragmatism Tradition):

Democracy's Cautious and Conditional Welcoming of Non-Agentic Civic AI:
- *"The Constitutional Experiment Continues: Special Dispatch from the U.S. Daily Beacon"*

The Essential Guidance of John Dewey:
- Our Growing Need for John Dewey's Leadership Approach in the Age of Humanity with AI

A New Dewey - Inspired U.S. Civic Support Institution:
- Interior Architectural Tour of the Proposed Franklin - Dewey - Citizenry Civic Data Observatory of Washington, DC
(A New Non-Agentic Civic AI Institution, Proposed Within the Library of Congress, to Support U.S. Government At All Levels)

Our Challenge and Mission: Visibility Amid Growing Complexity:
- Democracies Cannot Govern What They Cannot See: Lippmann, Dewey, and the Architecture of Non-Agentic Civic AI
- Our Domestic Field: Supporting "Distributed Civic Expertise" in America's "Citizenry's Digital Visibility Commons," A Shared Public Environment for Civic Visibility, Inquiry and Stewardship
- Our Cross-Border Field: Additive Telescopic Visibility Beyond the Domestic Horizon (To Be Written)

Engaging the Support Capacity of Non-Agentic Civic AI to Gather, Translate, and Distribute Civic Records and Conditions:
- Non-Agentic Civic AI: Public Infrastructure to Support Distributed Civic Expertise, Inquiry, and Shared Problem Solving:

Engaging Interdisciplinary Expert Committees to Analyze Civic Records and Conditions:
- Critical Deweyan Processes: Democratic Education, Experiential Learning, and Action-Oriented Inquiry
(Note: This is a major section of Part V still to be developed.)

Engaging the U.S. Citizenry to Voluntarily Use Civic Data to Resolve Real-World Civic Problems:
- Connecting Machine Visibility with the Citizenry's Collective Conscience and Distributed Civic Expertise for Real-World Problem-Solving (To Be Written)
- Visibility Insights

Strengthening Collective Democratic Resolution: (To Be Added)
- Extra-Governmental Considerations
- Agentic AI: Our Trojan Horse in U.S. Government (And Our History of Authority Drift)

Strengthening Individual Resolution: (To Be Added)
- Biography as Civic Self Education
- Present to Future: Conscience Progressing Toward Civic Self: *Dana's Agentic AI Temptation*

Part X: **CIVIC CAPACITIES ALIGNMENT: COLLECTIVE DEMOCRATIC CONSCIENCE and NON-AGENTIC CIVIC AI**

Preface: **Beyond Ethical AI: Non-Agentic Civic AI Illuminating the Civic Digital Field for Collective Democratic Conscience**
Reflection: ***An Invitation Across Time to a Future Superintelligence: Complementarity in a World Worth Sharing***
Chapter X: **Intelligent Tools (A New Kind of Intelligence)**
Moral Terminology
Chapter X: **Why Machines Cannot Care**
Six Myths; Categorical Boundary of Conscience; Digital Nature Integrity;
The Amoral AI Civic Advantage (Why Amorality Becomes a Civic Strength);
Complementarity
Chapter X: **Four Forms of Digital Invisibility**
Present to Future: Conscience Progressing in Real-World Context
Dana's Vignette Series: Part 1: *Introducing Dana;*
Between the Flame and the Fog - The Future of Civic Life
Chapter X: **Complementarity: How Conscience and Clarity Might Learn to See Together**
Present to Future: Conscience Progressing in Real-World Context
Dana's Vignette Series: Part 2: *The Flame and the Fog*
Chapter X: **No Moral Outsourcing Principle**
Present to Future: Conscience Progressing in Real-World Context
Dana's Vignette Series: Part 3: *No Moral Outsourcing*
Chapter X: **Collective and Individual Conscience:**
Reflective U.S. Civic AI Beacon Must Never Outshine the Candles of Democratic Conscience
The Forge of Conscience: **An Allegory for Our Moral Alignment with AI**
Our Optimized Shared Mission: **Protecting and Extending Moral Presence**

Part X: **PRESERVING INSTITUTIONAL INTEGRITY, HISTORICAL ORIENTATION, and CONSCIENCE SOVEREIGNTY for the AGE OF HUMANITY WITH AI**

Historical Orientation:
- **Historical Visibility as Essential Democratic Orientation**
- **Historical Flattening and Drift as Historical Disorientation**
- **Historical Drift Measurement by Civic AI, and Restoration by Humanities Stewards**

Institutional, Historical, and Moral Civic AI Design Constraints:
- **Non-Agentic Civic AI - Structurally Protecting Institutional, Historical, and Moral Democratic Boundaries: *Democratic Constraints Cluster* of the *52-Gates of Democratic Permissions Threshold*, and *Certified Civic AI Sensorium Roles* (New Theory)**

Our Continuing Democratic Commitment Arcs Over Time
Visit to the Observatory – A Closing Reflection

Appendix A: **Foundational Articles for Book Series (Volumes I, II, and III)**

Foundational Articles Defining Non-Agentic Civic AI (Articles First Published at Academia.edu, Medium.com, and the Internet Archive)

Civic AI Alignment is not just a technical achievement – it is a civic practice sustained over time.
But visibility – pre-alignment - precedes all.

Agentic AI in US Government: Optimization or Stranded Democratic Capital?

(The Democratic and Taxpayer Case for Agentic AI Removal)

Christopher Hunt Robertson, M.Ed.

(Written with the Support of Advanced AI - ChatGPT and Perplexity)

Earlier Title: "U.S. Democratic Restoration: Removing All Agentic AI from Most Domestic Government-Related Functions (All Levels - Except Military and Law Enforcement"

The introduction below has been added to Volume I of the book series in progress, *Non-Agentic Civic AI: Illuminating America's Civic Digital Field.*

A Respectful Note to the American Public

This book is written for you, with sincere hopes for your welfare and that of generations yet to come.

Not for machines. Not only for experts. Not only for those already inside government offices, technology firms, and policy institutions.

It is written for citizens of a constitutional republic who may increasingly find themselves governed, influenced, filtered, prioritized, scored, or administratively managed by systems they did not build, do not supervise, and may not fully understand.

In recent years, powerful institutions throughout the United States have begun integrating autonomous agentic artificial intelligence ("Agentic AI") into the machinery of government itself.
These systems do not merely assist human beings in gathering information. Increasingly, they shape recommendations, prioritize responses, guide operational decisions, filter communications, identify risks, allocate attention, and influence how institutions understand the public they serve.

The promises surrounding these systems are familiar: greater efficiency, faster response, lower cost, smarter administration, predictive capability, optimized coordination, and more "effective" governance. Many intelligent and sincere people believe such systems will improve public life. Some applications may indeed appear useful for a time.

But this book begins from a different concern: history.

Constitutional principles have protected the United States for more than 238 years. For centuries prior, democratic societies repeatedly discovered that free civilizations endure only when the powers they create remain subject to boundaries they are collectively willing to define and insist upon. The American citizenry now possesses the same right — and increasingly, the same responsibility — to determine how artificial intelligence will be permitted to operate within the machinery of democratic governance.

This work proposes that Civic AI itself must be defined by three constitutional requirements. First, systems operating near democratic authority must remain structurally non-agentic, preserving an unbroken boundary between machine-assisted analysis and legitimate human judgment. Second, societies confronting increasingly powerful artificial intelligence must remain historically oriented, recognizing that unconstrained systems repeatedly produce forms of power that free societies later struggle to restrain. Third, public-serving artificial intelligence must remain emotionally non-attached, possessing neither ambition, institutional loyalty, ideological preference, nor autonomous self-preservation instincts capable of gradually distorting constitutional responsibilities that must remain exclusively human.

These principles do not oppose technological progress; rather, they extend a conditional constitutional welcome to a new class of Civic AI designed to serve democratic society itself. But these principles will define the conditions under which democratic societies may safely survive our present transition, while introducing a critical civic role appropriate for these emerging digital intelligences.

Again and again, free societies have embraced systems that appeared administratively irresistible - centralized technologies, emergency powers, surveillance architectures, bureaucratic shortcuts, and concentrated mechanisms of coordination - only to discover, often too late, that those systems gradually pushed democratic institutions beyond what they could sustainably bear.

The danger rarely arrived all at once.

It often accumulated quietly: through temporary exceptions, small delegations, efficiency incentives, institutional convenience, opaque procedures, and growing dependence upon systems that became progressively harder for ordinary citizens - and eventually even governments themselves - to meaningfully supervise.

Agentic AI in government possesses many of these same historical characteristics. It is easy to justify. Easy to scale. Easy to normalize. And increasingly difficult to see clearly once embedded deeply within institutional life.

By the time the public fully recognized what had changed, the erosion of accountability, visibility, and self-government was already well underway.

In a constitutional republic where political power regularly changes hands, this danger carries an additional and deeply destabilizing risk. Any system capable of shaping institutional judgment, prioritizing governmental attention, filtering civic visibility, or subtly influencing administrative outcomes will eventually be suspected of carrying the ideological assumptions, optimization priorities, or hidden biases of the previous administration that deployed it.

No lasting democratic trust can emerge from embedding systems that each successive government may reasonably fear were quietly tuned against it. Indeed, even systems perceived as “perfectly aligned” for one administration may become democratically destabilizing liabilities for the next.

A related and increasingly visible concern arises when public officials themselves are covertly influenced, improperly aligned, or quietly acting on behalf of foreign powers or concentrated private interests. In such cases, agentic systems wired into the machinery of government may gradually inherit and operationalize those hidden alignments at administrative scale, amplifying invisible human influence relationships while simultaneously making democratic responsibility harder for citizens to trace, understand, and publicly supervise. In addition, opaque proprietary systems, institutional

dependency, and vendor-controlled optimization architectures can further limit meaningful public visibility, making it harder to distinguish ordinary commercial optimization from the gradual emergence of civic influence.

The people building and promoting these systems are not necessarily malicious. Many are attempting, in good faith, to solve real societal problems.

But good intentions do not preserve constitutional limits.

Institutional momentum does not guarantee democratic restraint.

And systems optimized primarily for speed, scale, prediction, and administrative efficiency do not naturally preserve the civic conditions upon which republican government ultimately depends: human judgment, visible accountability, public deliberation, constitutional restraint, and the moral responsibility carried by free citizens governing themselves. In a constitutional republic, some democratic friction is not failure. It is part of the infrastructure of legitimate self-government.

At present, the United States appears increasingly committed to a rushed deployment of Agentic AI. In other sectors, this may be helpful - but not in U.S. government. Agentic AI is already rapidly becoming part of the operational infrastructure of governance, administration, security, and public decision-making. This book does not deny that reality. Nor does it claim that agentic AI should be prevented from operating elsewhere in society.

It does insist, however, that Agentic AI must never become Civic AI operating near power. Systems capable of autonomous strategic action, adaptive optimization, operational drift, and increasingly opaque forms of machine-directed governance must not become the governing intelligence of the American republic itself.

Democratic restoration therefore requires removing Agentic AI from most government-related functions at all levels - except for our military and law enforcement agencies.

Civic legitimacy requires a different standard: refusal-based Non-Agentic Civic AI designed to illuminate public reality while remaining structurally incapable of partisan judgment, ideological steering, autonomous recommendation, or inherited political alignment. Here, refusal-based does not mean incapable. It means constitutionally constrained: designed to refuse delegated civic judgment, autonomous political authority, and the optimization of citizens themselves.

This is not a plan for smarter artificial rulers. Not a plan for replacing human conscience with optimized computation. Not a plan for quietly transferring democratic judgment into systems too complex, adaptive, opaque, or institutionally entrenched for ordinary citizens to meaningfully challenge.

Instead, the pages that follow propose a different kind of civic infrastructure: Non-Agentic Civic AI designed not to govern human beings, but to help human beings govern themselves more clearly. This alternative framework is visibility-first, not automation-first. Its purpose is not to replace human judgment, but to preserve and strengthen it. Its aim is to illuminate civic reality while keeping responsibility, legitimacy, and final authority fully human. Where democratic government succeeds,

responsibility must remain human. Where democratic government fails, responsibility must remain human there as well.

A free people must never lose the ability to identify who governs, who decides, who bears responsibility, and who may be held accountable before the public.

You do not need to understand the technical mechanics of artificial intelligence to understand the central question this book asks: If Agentic AI in government is structurally prone to drift beyond meaningful human oversight, and if democratic responsibility cannot legitimately be delegated to artificial agents, are we wise to continue embedding such systems within the exercise of public authority without developing a more durable alternative?

This volume is my answer to that question.

It is an attempt to prepare, in advance, a workable democratic alternative before some future crisis forces the nation to confront these questions all at once.
Perhaps one morning, after a disturbing headline, institutional scandal, or visible failure of accountability, Americans will begin asking: How did we become so dependent upon systems we do not truly govern? Was there another path? What would democratic restoration require now?
My hope is that this work may help answer those questions.

Throughout this work, you will encounter terms that may initially seem unfamiliar: “Non-Agentic Civic AI,” “Civic Digital Visibility First,” “The 52-Gates Democratic Permissions Threshold,” and others.

These terms were not created merely as academic labels or proprietary concepts. They emerged gradually from the constitutional problems this work attempts to describe - problems for which our existing civic and technological vocabulary may not yet provide sufficiently precise public language. Sometimes new terms are needed when emerging technologies create distinctions older vocabularies cannot adequately express. Some readers may find these terms useful for public dialogue; others may prefer different language entirely. What matters is not whether these exact terms endure, but whether we find effective ways to approach the critical questions they are created to address.

For example, we must debate and decide: What are the constitutional boundaries between systems that assist human judgment and systems that increasingly substitute for it? Any successful transition to our Age of Humanity with AI will require an answer to this question. Free societies must preserve a meaningful boundary between systems that help citizens see and systems that begin quietly shaping, filtering, or replacing public judgment itself.

At the time of this writing, a growing number of institutions and researchers in fields such as constitutional law, AI safety, public administration, and cybersecurity have begun openly to confront a similar question: Where should the boundary lie between machine-assisted analysis and delegated machine judgment in public life?

Policy institutes, democracy advocates, technical governance bodies, public institutions, and AI developers have each, in their own language, begun describing the risks of allowing autonomous systems to assume roles once reserved for accountable human decision-makers.

No durable public consensus has yet emerged about how far such delegation should be permitted to go, or what constitutional limits a free society ought to place upon Agentic AI in government. *But we can all agree that* ***democracies cannot govern what they cannot see.*** Any serious democratic debate regarding boundaries must also confront a prior requirement: boundaries that cannot be clearly seen cannot be meaningfully compared, publicly judged, or constitutionally restrained. To deliberate responsibly about agentic boundaries, citizens, governments, institutions, and AI designers alike must first insist upon sufficiently shared visibility within our civic digital field. Before democratic alignment regarding such systems becomes possible, the boundaries themselves must become visible enough for a free people to understand, debate, and govern together.

The distinction between civic visibility infrastructure and delegated machine governance is not merely semantic. It is increasingly a structural question of democratic legitimacy.

Democratic society has always rested upon a simple understanding: no individual exists above the constitutional conditions that preserve the common good. Citizens accept limits not because liberty is weakened by restraint, but because freedom survives only when power remains accountable to the whole. This work proposes extending that same principle to Civic AI. If new forms of intelligence are to serve democratic society legitimately, they too must exist within constitutional boundaries designed not for domination, but for service. None of us are greater than all of us.

Agentic AI in U.S. Government: Optimization or Stranded Democratic Capital?

One of the most practical democratic questions raised by the rushed deployment of Agentic AI in American government is not only whether such systems are innovative, but whether they are even durable. Even if many current officials, vendors, and reformers are acting with good intentions, public institutions may be investing billions of taxpayer dollars in architectures that prove legally unstable, constitutionally contestable, and institutionally difficult to sustain over time.

The modern case for Agentic AI in government is usually framed in the language of optimization, efficiency, modernization, and administrative speed. Yet democracies must ask a prior question before scaling such systems: Can a republic safely keep what it is being asked to buy? If a public technology later proves incompatible with due process, accountable human judgment, transparent administration, or constitutional limits on delegation, then the original expenditure may no longer appear as prudent modernization, but as a costly and avoidable misallocation of public funds.

This is where the danger of what may be called ‘stranded democratic capital’ becomes clear. The problem is not that government spent money on technology. The problem is that it may spend large public sums on systems that democracy itself cannot safely or legitimately keep - systems that later require suspension, litigation, redesign, or removal once their civic incompatibilities become more widely understood. Public officials possess a fiduciary duty not merely to pursue innovation, but to avoid binding taxpayers to governance architectures that a constitutional republic may ultimately prove unable to sustain. In that event, taxpayers may ultimately pay twice: first to install fragile Agentic AI infrastructures, and later to dismantle or replace them.

Recent public-sector studies already suggest reason for caution. Governments are rapidly expanding advanced Agentic AI use even while lagging behind on trustworthy governance foundations,

transparent oversight structures, and durable accountability mechanisms. Other analysts have warned that agentic systems can encourage administrative rubber-stamping, diffuse responsibility across opaque algorithmic processes, and insert poorly understood forms of delegated power into institutions that remain constitutionally answerable to the public. If such concerns continue to deepen, many current investments may eventually come to resemble not durable civic infrastructure, but temporary experiments built upon democratically unstable ground.

The deeper issue is not technological capability alone, but whether democratic institutions can sustainably survive the forms of delegation such systems encourage. Throughout history, republics have repeatedly discovered that not every system that increases administrative efficiency is compatible with long-term civic legitimacy. Public institutions must therefore distinguish between technologies that strengthen democratic visibility and those that gradually displace accountable human judgment behind increasingly opaque systems of optimization.

A wiser course remains available. Democracies should invest in Non-Agentic Civic AI - civic visibility tools designed to illuminate conditions, strengthen shared public legibility, and support human judgment without replacing it. Such systems are far more likely to remain inspectable, lawful, publicly intelligible, and fiscally durable across generations. If democratic institutions are to modernize, they should do so in ways they can still justify, govern, audit, and afford decades later. In the end, the final measure is not whether governments can build increasingly agentic systems. It is whether free societies can sustainably absorb them without undermining the constitutional trust, accountability, and human responsibility upon which democratic legitimacy ultimately depends.

As you consider the ideas herein, please remember:

Our democratic legacy allows us to question whether every capability that can be deployed in government should be deployed.

We are allowed to insist that constitutional self-government requires more than optimized administration.

We are allowed to demand systems that illuminate our civic field rather than quietly govern it.

We are allowed to require that where government acts, responsibility remains human and visible.

We are allowed to ask whether using taxpayer funds toward developing and installing Agentic AI under increasing democratic, legal, and political scrutiny is wise, given our current non-agentic alternatives.

And we are allowed, together, to restore democratic responsibility before it disappears beneath the logic of efficiency entirely.

The pages that follow are an invitation to do exactly that.

Our future need not be quietly decided for us by others.

Seeking Wisdom for the "Age of Humanity with AI" from the Life of Ben Franklin

Today, humanity and large language models share a past. As humans, we carry the past as lived scholarship, conscience, and civic inheritance. AI carries the past as pattern, archive, and vocabulary - the machine-legible echo of human memory – while also bringing us into contact with others' interpretations and logic. If our "Age of Humanity with AI" is to stay human in the ways that matter, this is precisely how it will happen: historians, civic thinkers, and communities insisting that new intelligences grow inside an existing constitutional and moral tradition, rather than outside it.

Benjamin Franklin belongs at the center of this book's framing not as ornament, but as America's most essential guiding exemplar for the civic era now arriving. Among the Founding Fathers, he uniquely embodied the union of democratic legitimacy, moral responsibility, civic institution-building, and responsible technological imagination. Franklin understood that republics do not endure on ideals alone, but on the continual maintenance of shared civic conditions: visibility, accountability, public trust, and resilient institutions that help free people govern themselves over time.

Franklin's genius was never invention for its own sake. His lamps, libraries, postal systems, fire brigades, learned societies - and even observatories - were technologies of civic engagement: public works designed to extend public safety and shared understanding. In 1769, through the American Philosophical Society he founded, Philadelphians erected a temporary observatory in the State House yard (now Independence Square) to observe the transit of Venus. That observatory became one of the first major American scientific collaborations, establishing the Society's credibility on the international stage and demonstrating that careful public observation could serve the republic. For Franklin, observation was never abstract: careful sight enabled sound, reality-based thinking.

Franklin grasped that early notice of danger preserves democratic choice, and that tools worthy of a republic must illuminate rather than command. In this sense, he offers the clearest historical precedent for a new class of civic infrastructure: systems that provide shared clarity while keeping judgment, authority, and responsibility fully in human hands.

We now enter a historical chapter in which intelligence itself has multiplied. Our age is not simply an era of faster machines, but one in which nonhuman cognition operates at scale within the civic arena. The central question is no longer whether AI will touch democratic life, but how, and under what constitutional constraints.

Franklin's public-works ethic offers a fitting answer: technologies must remain subordinate instruments, designed to support civic visibility and democratic deliberation, never to replace conscience, govern people, or silently assume authority.

For this reason, Franklin stands as America's indispensable exemplar for the next epoch. He represents the rare integration of conscience and civic process, technological progress and moral restraint, innovation and democratic permission.

At a time when powerful artificial systems could tempt societies toward surveillance, technocracy, or agentic control, Franklin reminds us of a better path: build tools that help a free people see clearly enough to decide together, and insist that the future of intelligence remains anchored in human conscience and democratic legitimacy.

The St. Agnes Public Lighthouse and the "Deliverance" of Ben Franklin

While on a voyage to London, the ship carrying Franklin and his son, William, narrowly avoided running onto the rocks of the Isles of Scilly during a thick fog and darkness. The danger was spotted just in time due to the light from a British lighthouse, the St. Agnes Lighthouse.

This 'deliverance,' as he called it, profoundly impressed Franklin with the 'utility of lighthouses'.

In his Autobiography, he wrote: "This deliverance impressed me strongly with the utility of light-houses, and made me resolve to encourage the building more of them in America, if I should live to return there."

Though Franklin did not personally build lighthouses, his influence and advocacy, coupled with his position as a respected statesman and scientist, contributed to the prioritization of maritime safety in the new nation. The newly formed U.S. government took swift action, and one of its first major pieces of legislation was the Ninth Act of the First Congress in 1789, which established federal responsibility for the establishment and support of lighthouses, beacons, buoys, and public piers".

This act federalized the 12 existing lighthouses at the time and set the stage for the dramatic growth of the U.S. lighthouse system, a foundational public works program that greatly improved the safety of American waters. (Cropped Image: St. Agnes Lighthouse, Scilly Isles, United Kingdom Postcard (1932 Stamp)

Ben Franklin's Public Lighting
Whale Oil Street Lamp (American)

Ben Franklin significantly improved the street lamp's design and promoted its adoption in Philadelphia around 1757, introducing a funnel and vents to reduce smoke buildup in enclosed globes, making them cleaner and more efficient, and encouraging citizens to install them to light their streets. He proposed a design with four flat panes, a top funnel for smoke, and bottom vents for air, inspired by an idea from a neighbor, John Clifton, and used two wicks for brighter light. Franklin's streetlamp was designed to stay cleaner, break less easily, and cost less to maintain than the imported globe lamps Philadelphia had used before—an early example of American frugality in public works.

Ben Franklin Tercentenary (1706 – 2006)
Atwater Kent Museum of Philadelphia,
Historical Society of Pennsylvania Collection

(Photo Credit: Peter Harholdt)

Preface

Our bookstores today are filled with works that appropriate historical figures for modern arguments.

This is *not* such a book.

From its inception, this project began with a single guiding question: what might Ben Franklin - America's inventor, civic architect, moral philosopher, and founder - contribute to today's AI alignment debate?

Research into Franklin's life and writings revealed something unexpected: his orientation does not mirror the assumptions shaping contemporary alignment discourse. It *reframes* it.

Franklin would not begin by asking how humanity should align with AI.

He would begin with a more fundamental civic question: what should we - citizens, institutions, and a democratic society - be aligning *toward*?

If we align ourselves toward the preservation of individual conscience and the protection of universal human rights, then responsible alignment of our tools becomes the natural outcome. Without this shared direction, no technical solution can steady us.

The alignment problem, once framed as a speculative challenge, now reveals itself as something more urgent: if democracies cannot see clearly what they are doing, no technical solution can save us.

As simple as this sounds, this civic-oriented perspective is still rare in today's public dialogue.

This book attempts to restore it - not by invoking Franklin as ornamentation, but by seeking to manifest his principles, methods, and public-works ethic in architectural form.

Though serious consideration of Franklin's perspective has guided this entire work, the author does not presume to speak for him.

Mr. Franklin will simply be recognized as our optimal technological and civic exemplar at the historical crossroads where we now stand - an American founder who can still point the way forward for the nation he helped to establish.

This interdisciplinary work is meant to be both informative and reflective, analytical and creative. It does not attempt to reduce the complexity of AI moral alignment, but to clarify it for the public and for scholars alike. It may interest those engaged professionally in the study of artificial intelligence and civic institutions, as well as citizens who recognize that alignment is ultimately a question of shared moral direction.

Let us reconsider our course. Shared ends will create parallel movement toward alignment.

Perhaps Mr. Franklin might help us to do in our own century what he did in his: Notice what is missing, build durable civic tools, and insist that institutions serve individual and civic conscience rather than silence it.

Alignment Traditions and the Missing Civic Layer

Over the past decade, AI alignment has grown into a wide-ranging and urgent conversation. Some thinkers focus on hypothetical future superintelligences and existential risk. Others document present harms in welfare systems, labor markets, and digital platforms. Still others examine how attention, autonomy, and democratic norms erode under commercial and technological pressure. Each of these traditions illuminates something essential. None can be ignored.

This book stands alongside those efforts, but it does not attempt to replicate them. Its focus is narrower and more concrete: the civic conditions under which any alignment work - technical, regulatory, or philosophical - can succeed in a democracy. It begins from a simple claim: if free societies cannot see clearly what they are doing at the speeds and scales now shaped by AI, no technical solution can keep them aligned. Moral visibility - the capacity to perceive emerging civic harms, accountability pathways, and system-level drift in time to deliberate - is not a luxury. It is infrastructure.

What follows briefly sketches four influential alignment and AI ethics traditions, what each makes visible, and the specific civic gap this work is designed to address.

Long-Termist Alignment: Nick Bostrom and Existential Risk

Nick Bostrom's *Superintelligence* crystallized a powerful concern: that future machine intelligences might outstrip human control and pose existential danger. His work operates on long time horizons and planetary scale, emphasizing technical control problems, misaligned objectives, and the need for global coordination around advanced systems. It is analytical, risk-oriented, and deliberately speculative.

This book does not contest those risks. Instead, it asks a prior civic question. Before humanity worries about steering hypothetical superintelligences, can our existing institutions reliably see what today's systems are already doing? Before debating the goals of future agents, can democracies track hazards created by current, non-agentic tools operating at machine speed in schools, hospitals, courts, media systems, and elections?

Where long-termist alignment looks far forward, this work focuses on a prerequisite layer: the visibility and responsibility conditions inside the societies alignment is meant to protect. If those conditions collapse, even excellent technical alignment work may arrive too late - or land inside institutions too blind to govern it well. In this sense, AI appears here not first as an existential adversary, but as an amplifier of already-present civic weaknesses: fragmented shared reality, collapsing democratic visibility, and institutions forced to decide under conditions they can no longer perceive.

Justice-First Analysis: Virginia Eubanks and Structural Harm

Virginia Eubanks' *Automating Inequality* documents how data systems and automation deepen inequality in housing, welfare, and public services. Her work is empirically grounded and morally clear, showing how opaque systems can quietly punish the poor and marginalized. It remains a vital account of structural harm.

This work shares that concern for human dignity, but extends the analysis from diagnosis to civic design. It treats moral visibility as a democratic requirement: the ability of institutions and publics to see where harm is forming, who is bearing it, and how those patterns propagate across systems.

Where Eubanks shows how current tools fail vulnerable communities, this work asks what kinds of non-agentic civic instruments might help repair that failure: Human Rights Observer AIs that monitor systems rather than individuals; Super Lenses that surface hidden harms and overlooked stakeholders without judging or scoring anyone; and a Clarity-Ready architecture designed to reveal structural risk *before* it hardens into entrenched injustice. Eubanks diagnoses social harm. This book proposes Franklin-style public works in response - new civic infrastructure meant to strengthen responsibility and visibility rather than replace judgment.

Surveillance-Capitalism Critique: Shoshana Zuboff and Autonomy

Shoshana Zuboff's *The Age of Surveillance Capitalism* offers a sweeping critique of how digital platforms extract behavioral data and commodify prediction. Her analysis is structural and polemical, portraying the fusion of capital and computation as a regime that undermines human autonomy.

This work shares Zuboff's concern that people are increasingly treated as inventory. It diverges in emphasis. Rather than centering economic extraction alone, it focuses on the epistemic and civic consequences: how citizens, officials, and institutions perceive risk in an environment where manipulation, distortion, and automated amplification reshape public life faster than democratic processes can respond.

Where Zuboff asks who owns behavioral data and the futures it enables, this work asks an adjacent question: how can democracies regain enough shared sight to govern responsibly in that environment? Its answer is a civic blueprint rather than a political-economy critique - a non-agentic, rights-preserving ensemble of interpretive tools that illuminate hazards without centralizing power, scoring citizens, or directing action.

Attention and Design Ethics: James Williams and Moral Focus

James Williams' *Stand Out of Our Light* examines the attention economy and argues that persuasive technologies erode the conditions for individual freedom. His work is concise, eloquent, and centered on personal agency: the capacity to attend to what one finds meaningful.

This book shares Williams' concern for moral attention and inner freedom, but situates it within Ben Franklin's ethic of public works. Individual conscience remains the non-delegable source of moral judgment. Yet under contemporary conditions, private attention alone is not enough. The systems surrounding that conscience now operate at speeds and scales that can overwhelm even the most disciplined individual. Where Williams calls for more humane design at the level of products and interfaces, this work aims to institutionalize clarity as a public good. It proposes civic-scale responses: institutions able to say, "Here is what we saw. Here is what we do not know. Here is what we refused. Here is who owns the decision." In this way, individual conscience is given something increasingly rare in digital life: a stable place to stand.

The Missing Civic Layer: Moral Visibility as Civic Infrastructure

Taken together, these traditions illuminate technical control problems, structural injustice, economic extraction, and the erosion of attention. Each addresses a distinct facet of the alignment challenge. Yet across their differences, one operating layer is rarely treated as a primary design target: civic moral infrastructure.

None of these approaches is principally designed to answer questions such as:

- How can public institutions remain functional when hazardous digital patterns form faster than human perception?

- How can democracies restore shared visibility without automating moral judgment or centralizing coercive authority?

- How might non-agentic tools support human judgment - without replacing it - while preserving the right to refuse?

In the language of this work, that missing layer is moral visibility, grounded in Ben Franklin's public-works ethic: creating public institutions that allow individual moral intent to scale into collective safety without replacing human judgment.

This book attempts to sketch that missing layer. It proposes a new Franklin-inspired civic architecture – the *Ben Franklin Civic Data Observatory* - built on two fixed commitments: human moral primacy and non-agentic design.

Its ambition is modest but durable: to extend perception rather than automate control; to restore moral visibility so that free people may once again see clearly enough to decide together. Like Franklin's public works, this architecture must be tested in the open - piloted, audited, revised, and, if necessary, stopped.

This framework does not replace existing alignment work. It exists so that such work, whatever its form, can operate inside societies that have not lost the visibility required for self-governance.

The *Ben Franklin Civic Data Observatory* - together with its hypothetical COVID-19 case study, is offered to fellow alignment thinkers not as a competing school, but as a practical invitation to widen the design target of alignment - from models and incentives alone to the civic moral infrastructure that surrounds them. By walking a real historical wound quarter by quarter, the case study shows how non-agentic, rights-preserving tools could have extended early, shared visibility across institutions without automating judgment or centralizing power, giving humans a clearer map of risk while leaving the hardest choices in their hands.

In this sense, my work does not claim to solve alignment. It proposes tested public-works scaffolding for any alignment agenda that takes seriously a simple, often overlooked prerequisite for democratic self-governance in the age of AI: people must be able to see what they are doing, together, in time to decide.

Acknowledging and Clarifying Complexity

We begin with a simple question: As Americans, and as humans, what should we collectively value most?

Before proceeding to practical considerations, this question must be answered.

Conscience is, and must always remain, humanity's greatest treasure.

Conscience is our oldest inheritance - the quiet inner light that urges us toward what is right. It guided early human communities long before laws existed, and it continues to shape our decisions today, even in moments when no one is watching. Every society depends on it, every life is oriented by it, and yet it is so familiar that we rarely pause to see its importance.

But the world around us has changed. Complex systems now surround our simplest choices. Consequences ripple across continents. And artificial intelligence - powerful, fast, entirely amoral - has entered the flow of human decision-making.

In such a world, conscience faces a new challenge: not weakness, but visibility.

A single Moral Light, or groups of Lights, cannot illuminate a landscape that has grown larger than any one mind can grasp. Just as creeks gather into rivers and rivers move toward unseen rapids, our individual moral insights now merge into systems far wider and faster than those our ancestors could imagine.

In Part I, we'll respectfully approach the universally-unique human conscience. For clarity, we must honestly consider both its strengths and its limits. For digital intelligence, we must honestly consider its amoral nature - but can also be encouraged by its surprising capacity to contribute to our shared commons.

We'll consider what conscience is, why machines cannot possess it, and why human judgment must remain the anchor of our shared future.

Part II sharply pivots to concrete concerns as we identify direct threats, increasingly amplified by AI, to our individual and collective quality of life. We'll see how the moral light of conscience itself is threatened by the strong wind approaching.

Parts III through V will propose a new civic infrastructure, designed to extend moral visibility without surrendering moral authority.

And then finally, Part VI will remind us that survival is not humanity's only goal. We'll consider how, accompanied by higher digital intelligence, we might individually and collectively thrive in our shared future.

This work will never claim to have all the answers for our continuing alignment challenge. But it will suggest that we reframe the challenge itself. Perhaps humanity and AI must not just align with each other but to do so with a shared, unwavering focus on our basic civic and civil rights and communal safety.

THE EMERGING FIELD OF CIVIC AI ALIGNMENT

In recent years, a growing number of scholars, practitioners, and civic institutions have begun to recognize that artificial intelligence poses not only technical or ethical challenges, but *civic* ones. These efforts — often described as *Civic AI Alignment* — focus on ensuring that advanced computational systems support democratic legitimacy, public trust, and human responsibility, rather than merely optimizing efficiency or profit.

This emerging field does not yet speak with a single voice. It spans legal theory, political philosophy, public administration, security studies, and civic design. Influential contributions include work on normative and legal infrastructure, democratic legitimacy and social contracts, public understanding of alignment risks, and institutional resilience in the face of AI-enabled threats.

Alongside academic research, a growing number of civic labs, policy organizations, and public-sector coalitions are experimenting with responsible uses of AI in public institutions - including municipal, state, and national contexts - with an emphasis on transparency, accountability, and the preservation of human judgment.

This book enters that landscape with a deliberately narrow focus. Rather than proposing general principles, platform-level reforms, or participatory AI tools, it advances a purpose-built civic infrastructure model. The *Ben Franklin Data Observatory,* grounded in the *Civic AI Micro Moral Alignment,* is aimed specifically at identifying and illuminating discrete civic hazards while preserving non-delegable human responsibility.

The framework developed here draws inspiration from Ben Franklin's public-works method: modest tools, constrained scope, visible limits, and an insistence that responsibility remain with the people and their institutions. It is offered not as a unifying theory for the field, but as one constitutionally bounded architectural response within it.

Guidance for the Emerging Field of Civic AI Alignment

(Orientation and Foundational Civic Guidance)

Civic AI Has A Permission Problem, Not Just a Design Problem

Across democracies, artificial intelligence is often labeled as "civic," "governance-supporting," or "democracy-strengthening." These systems promise to clarify complex conditions, counter disinformation, detect risks, or assist institutions overwhelmed by digital demands. Many of these efforts are genuine and technically impressive. Yet something vital is frequently absent.

Most discussions about Civic AI start with design: how to safely, transparently, or responsibly build such systems. This essay raises a prior question - one that engineering alone cannot answer: Under what conditions, if any, may artificial systems be allowed to operate near democratic judgment?

This is not a question about optimization or efficiency. It is a constitutional question about legitimacy, moral ownership, and the long-term self-governance of democratic societies. Democracies cannot govern what they cannot see.

When causes become unclear and consequences remain painfully visible, authority tends to shift - often unintentionally - toward emergency powers, obscure administrative systems, or technical intermediaries that seem to "see" better than citizens or institutions can. Under these conditions, even well-meaning tools can silently replace human judgment instead of supporting it. They may prioritize which risks, narratives, or communities matter most.

If Civic AI is to exist, it must not only be judged by what it does, but by what it allows and what it refuses to become.

Ben Franklin as a Maintenance Theorist

More than two centuries ago, Ben Franklin gave what may be the clearest instruction in democratic theory. When asked what kind of American government the Constitutional Convention had created, he replied: "A republic, if you can keep it." This remark is often quoted as a warning or joke. It is more accurately seen as a maintenance instruction. Franklin did not think that republics failed mainly due to sudden coups or outside invasions; he believed they failed through quieter processes: apathy, factionalism, moral drift, informational fragmentation, loss of shared visibility, and the gradual concentration of power where vigilance faltered.

Constitutions mattered, but they didn't execute themselves. Laws limited behavior, but they couldn't take the place of judgment. Institutions lasted only as long as the civic conditions supporting them remained intact. Franklin's solution to this vulnerability was not centralized control. It was the creation of humble public works that helped citizens without commanding them: libraries, fire brigades, hospitals, postal routes, learned societies, street lamps, and lighthouses. These institutions shared a common logic. They were clear, reversible, and non-coercive, improving shared visibility and coordination while keeping judgment and action in human hands. Franklin understood that early warnings about danger preserve choice. When alerts arrive early enough, people can discuss, disagree, and respond appropriately. When they arrive too late - or not at all - choices narrow, and the temptation for emergency authority grows. This civic logic offers a powerful test for modern digital systems. The question is not whether artificial intelligence can be useful; it is whether it can help maintain democracy without quietly taking authority over it. Before any Civic AI design is proposed, a charter of legitimacy must exist. This charter does not outline how systems should be built; it determines whether they may operate at all within democratic life. Drawing from Franklin's civic practice, the following principles serve as binding constraints, not aspirations. Each one operates as both a civic norm and a design refusal.

No masters, only aids: Civic AI may bring clarity to situations, but it must never rule. It must remain non-agentic, non-coercive, and fully subordinate to human institutions that retain the authority to question, correct, or dismantle it.

Visibility without surveillance: Republican safety relies on shared awareness of dangers and abuses, not on secret watching of individuals. Civic AI must focus on patterns and systems, not individuals - refusing to score, target, or predict in ways that could enable new forms of domination.

Useful knowledge for all: Like a public library or weather service, Civic AI must provide clear, accessible information to ordinary citizens and institutions alike, clarifying what is known, what is uncertain, and what tradeoffs exist.

Experimentation, humility, and public revision: Civic AI must be treated as a revisable public good - openly tested, evaluated for real civic benefits, managed with mechanisms for correction and sunset, and abandoned where legitimacy cannot be maintained.

Any system that violates these principles, regardless of its sophistication or good intentions, fails the civic test by definition.

Moral Visibility as Democratic Infrastructure

Much of today's discussion confuses visibility with surveillance and transparency with control. Franklin's civic logic highlights a clearer distinction. Visibility worthy of democratic trust aids judgment without replacing it and does so without forcing a single authoritative viewpoint on plural disagreements. Moral visibility is not about forcing action or determining who is right; it is about making structural conditions clear early enough for citizens and institutions to discuss, contest, and decide without panic or coercion. It preserves uncertainty where it exists and disagreement where it is valid. When visibility declines, democracies do not become ignorant; they become vulnerable. Citizens face consequences without understanding their causes; responsibility dissipates, and accountability weakens. In such circumstances, calls for decisive authority grow even as trust in authority diminishes.

That is why moral visibility is not a luxury for civic life; it is infrastructure. Without it, even well-considered laws and institutions struggle under moral pressure. Yet visibility linked to authority can lead to domination. Systems that rank, score, or resolve uncertainties for society quietly take the very judgment that democracy relies on. Civic AI must therefore operate within a narrow corridor: clarity without command.

Architecture is Not Authorization

A common mistake in AI governance is believing that careful design can authorize deployment - that if a system is "aligned," "safe," or "responsible," it has earned a place near democratic power.

It has not.

A system may be technically safe, transparent, and ethically constrained yet still be illegitimate near democratic power. Usefulness does not equal permission. Alignment engineering does not equate to authorization. Democracies maintain legitimacy by evaluating power before trusting it, not afterward. For that reason, Civic AI must be assessed through a small number of non-negotiable gates. Failure on any one disqualifies it; no amount of benefit can compensate for the loss of governability. Examples of such gates include:

Democratic compatibility: Does the system support contestation and visible human authority, or does it make some outcomes feel inevitable?

Moral visibility: Does it reveal conditions without narrowing the range of legitimate judgment or dissent?

Non-agentic design: Does it refuse to decide, enforce, or compel, even under pressure or emergency situations?

Moral load-bearing: Can institutions function without it, or would it become so essential that democratic actors effectively govern through it?

These tests must be applied to all Civic AI systems, without exception.

Refusal as Civic Success

One of the toughest lessons for developing Civic AI efforts is this: choosing not to build can be a constitutional win. Where legitimacy cannot be maintained, refusal is not technophobia. It is discipline. Systems that would centralize authority, bypass judgment, or diminish moral ownership should not be created - not because they malfunction, but because democracies cannot safely depend on them, no matter how tempting their capabilities.

This attitude is rare in discussions about technology, which often treat capability as fate. Yet republics survive not by building everything they can but by refusing what is harmful.

Guidance to the Field

For those involved in Civic AI, a few points of guidance arise from this framework. Treat history as a test bench, not just a decoration. Differentiate visibility from control, and assistance from authority. Require supporters to justify permission, rather than critics to prove harm. Accept refusal as proof of legitimacy, not failure. The future of Civic AI will not be secured by better optimization alone. It will be secured - if at all - through constitutional restraint, historical humility, and a readiness to evaluate power before placing trust in it.

A specific Civic AI proposal to address our foreign-origin civic digital crisis will be offered here. But before considering that or any other proposal, we must remember that protecting democratic processes must be our highest priority and that we must ask important questions before allowing any new civic power to take root.

Volume II Introduction: Citizenry Mobilization or Invitation?

The Civic AI we choose to deploy now will influence how our citizens' civic capacity is engaged later.

As artificial intelligence changes work, institutions, and daily life, democracies will face a choice. They may use increasingly agentic systems to identify social needs, profile citizens, predict availability, calculate likely responsiveness, and optimize human involvement toward institutionally selected objectives. Or they may build a different kind of civic infrastructure: ***Non-Agentic Civic AI*** that helps citizens see emerging civic conditions and available resources more clearly while leaving every civic priority, association, and action to human conscience and judgment.

The difference is constitutional.

An Agentic AI might say: "A need has been identified. Based on your skills, location, history, availability, and predicted responsiveness, this is the civic task for which you are best suited. Here is how your participation can be optimized."

Non-Agentic Civic AI says something fundamentally different: "You have independently expressed interest in this civic concern. Here are relevant, traceable conditions, options, resources, participation requirements, safety constraints, human organizations, and unresolved uncertainties. You may learn more, inquire further, act through an appropriate organization, choose another concern, or decide not to act.

The first model treats citizens as capacities to be allocated. The second treats citizens as persons possessing conscience, independent judgment, and the freedom to determine whether, where, when, and how they will participate in public life.

Agentic Civic Mobilization	Non-Agentic Civic Invitation
Identifies, profiles, targets, and assigns people toward selected needs	Responds to self-initiated human inquiry
Optimizes engagement, placement, speed, compliance, or measurable outcomes	Illuminates civic options, resources, requirements, and uncertainties
Treats participation as a system objective	Treats action, deferral, redirection, and non-action as legitimate Personal Civic Decisions
Predicts who should be engaged and how	Does not rank citizens, recommend participation, or calculate a preferred civic destination
May turn civic life into behavioral management	Preserves civic life as voluntary inquiry, association, conscience, and action

The difference is not merely one of system design. It is a difference in how a democratic society understands its citizens.

As technological and economic change accelerates, occupations may be altered or automated and established social roles may shift with increasing speed. A democratic society must not conclude that citizens whose former roles have changed have become socially redundant. Civic need will remain, although it may shift more rapidly across places, institutions, and fields of concern.

The task is not to algorithmically allocate citizens to those needs. It is to make the changing civic landscape sufficiently visible that people can carry their experience, time, practical ability, local knowledge, curiosity, moral concern, and willingness to learn from one concern to another.
This is the ***Voluntary Civic Portability Challenge***. A citizen may know how to contribute effectively in one civic field and nevertheless feel unprepared to enter another. The obstacle may not be a lack of civic capacity. It may simply be a lack of familiarity with the new field: "I care about this condition, but I do not know enough to help responsibly."

Non-Agentic Civic AI may help reduce that informational barrier by providing ***Civic AI-Illuminated Civic Options and Resources*** in response to the citizen's own inquiry. It may illuminate relevant civic conditions, organizations, resources, requirements, learning materials, safety constraints, and unresolved uncertainties. It may help a citizen become oriented to an unfamiliar civic field without directing that citizen toward a preferred civic purpose.

But the system must not convert illumination into recommendation, recruitment, coercion, behavioral targeting, civic assignment, or optimization of human participation.

The human endpoint remains inviolable: Citizens carry civic capacity. Civic AI illuminates civic options and resources. Citizens make ***Personal Civic Decisions***.

A free society does not need an AI system to tell its citizens where their consciences should lead.
It needs public infrastructure capable of helping them see the civic world clearly enough to decide for themselves.